

Segmentación de Imágenes Médicas mediante UNET

Steve Avendaño¹, Jimena Olveres^{2,3}, Boris Escalante-Ramírez^{2,3}

¹Facultad de Ciencias; ² Facultad de Ingeniería; ³Centro Virtual de Computacion. UNAM CDMX.



Las enfermedades cardiovasculares constituyen la primera causa de enfermedad y muerte en el mundo occidental, con un continuo avance en los países en vías de desarrollo. Esto a generado un aumento en la información e imágenes que se genera en el quehacer diario de la práctica clínica, lo que aumenta la importancia del desarrollo y uso de nuevas técnicas de automatización en esta área médica. En este trabajo mostramos una aplicación con aprendizaje profundo, utilizando la red U-NET para llevar a cabo la segmentación en el volumen cardíaco.

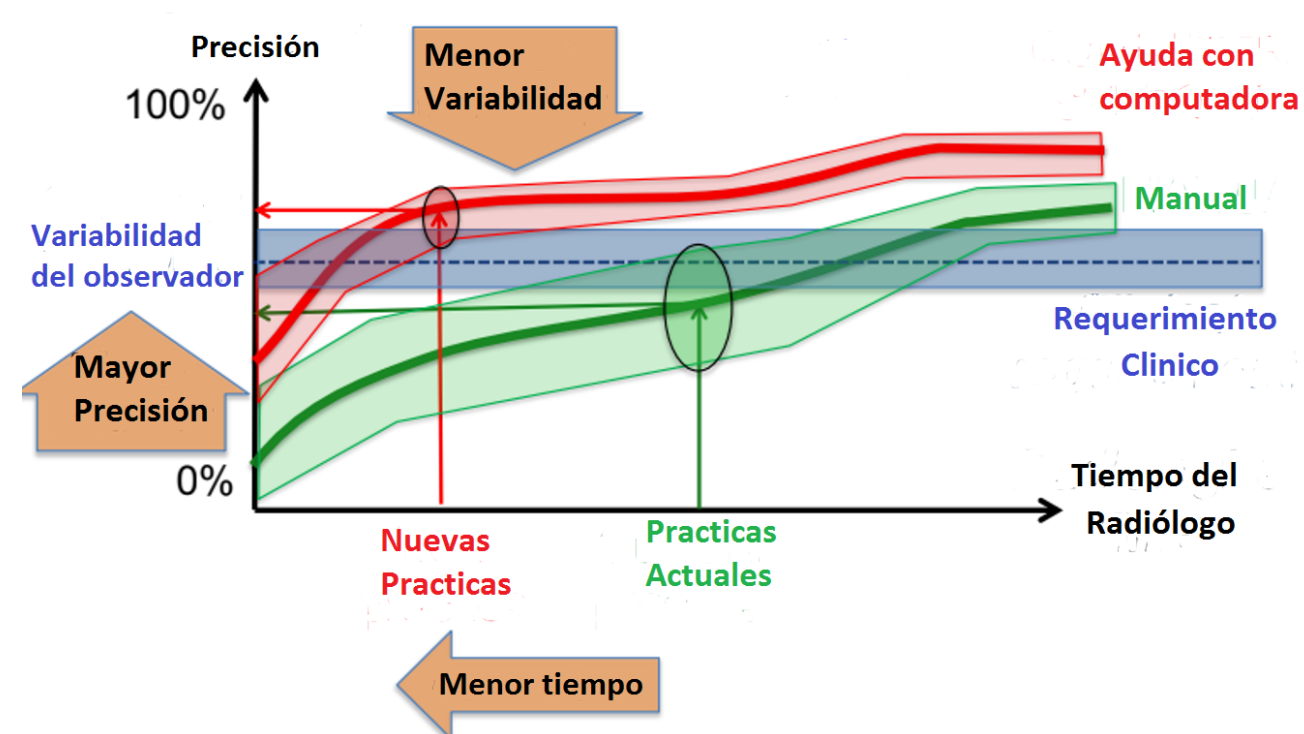


Fig1. Ventajas que presenta la automatización de procesos repetitivos, particularmente en el área de imagenología médica.

Selección de Modelo

Para la selección del modelo, se tomó en cuenta el trabajo llevado a cabo por Olaf et al.[4], el cual consiste en una arquitectura tipo autoencoder, con "skip connections", estas ultimas tienen la finalidad de disminuir la pérdida de la información espacial (localización) que ocurre durante el proceso de codificación y posterior decodificación, lo que mejora el proceso de segmentación.

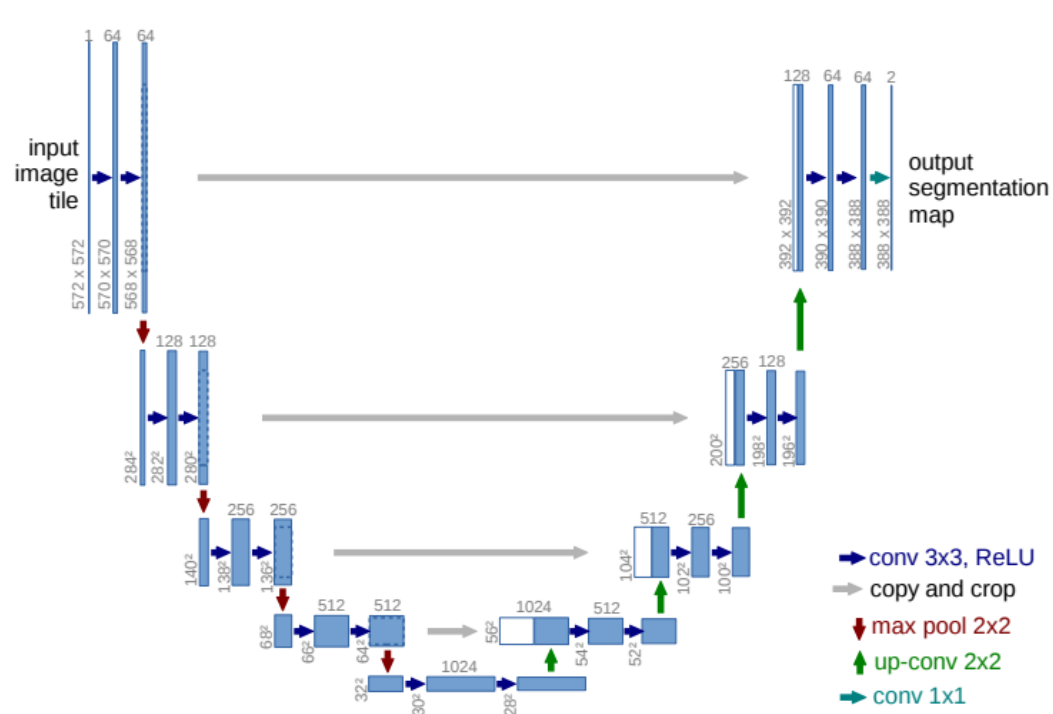


Fig2. Ejemplo de una arquitectura tipo U-NET. [4]

Las diferencia principal entre la arquitectura empleada en este trabajo y la original consiste en el uso de convoluciones separables, las cuales fueron introducidas por Laurent Sifre [6], esta operación permite desacoplar el proceso de aprendizaje de la correlación entre píxeles vecinos y entre los píxeles de distintos canales.

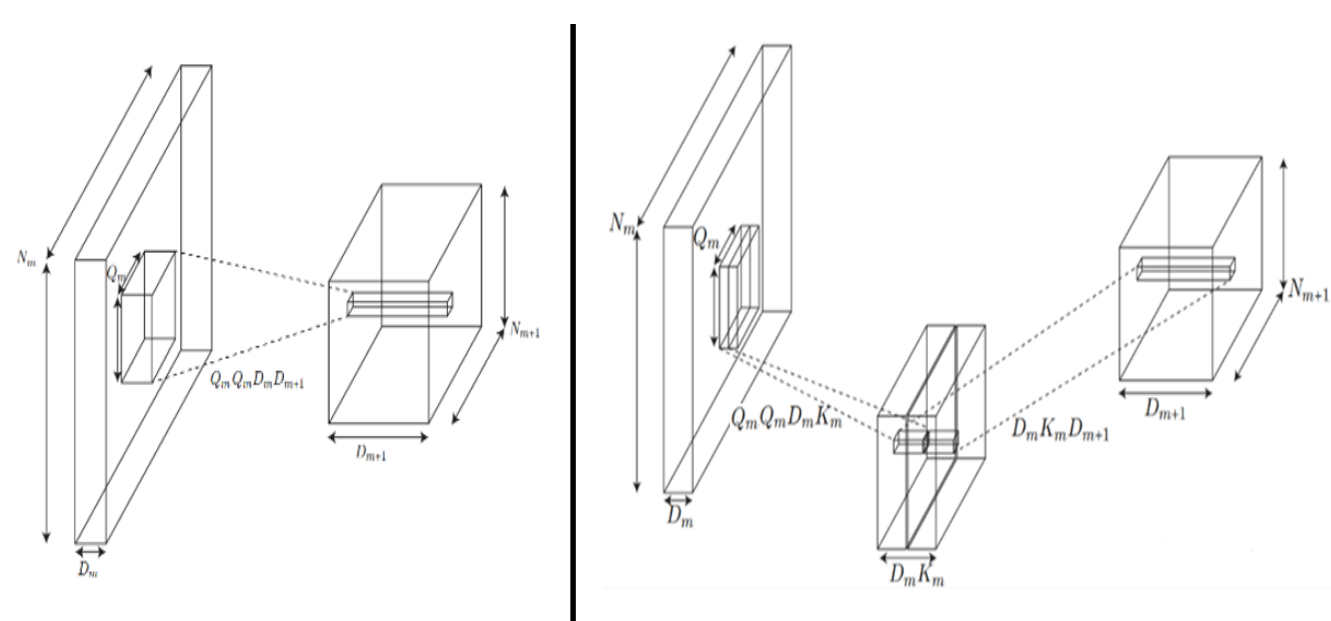


Fig3. Diferencia entre los operadores de convolución. Columna Izquierda) Convolución convencional. Columna Derecha) Convolución separable [6]

Conjunto de Datos

Las imágenes junto con sus etiquetas fueron tomadas de un conjunto de datos de uso libre, el cual consiste de imágenes adquiridas en un cierto volumen a nivel del tórax, estas

fueron adquiridas por Resonancia Magnética Cardiovascular durante una práctica clínica en el Hospital de Niños de Boston, Boston, MA, USA. Los casos incluyen una variedad de enfermedades congénitas del corazón. Algunos de los pacientes han sido sometidos a intervenciones [3].

Imágenes y Etiquetas

La clase a la que están asociados los volúmenes de sangre incluye las aurículas izquierda y derecha, los ventrículos izquierdo y derecho, la aorta, venas pulmonares, arterias pulmonares, así como la vena cava superior e inferior. La clase asociada al miocardio incluye la pared muscular que rodea los dos ventrículos y el septum que los separa. Se llevo a cabo un proceso de 'data augmentation' en el cual se aplicaron transformaciones como desplazamiento, rotación, magnificación y shearing a las imágenes y mascarar en conjunto. Estas nuevas imágenes eran posteriormente introducidas a la red para su entrenamiento.

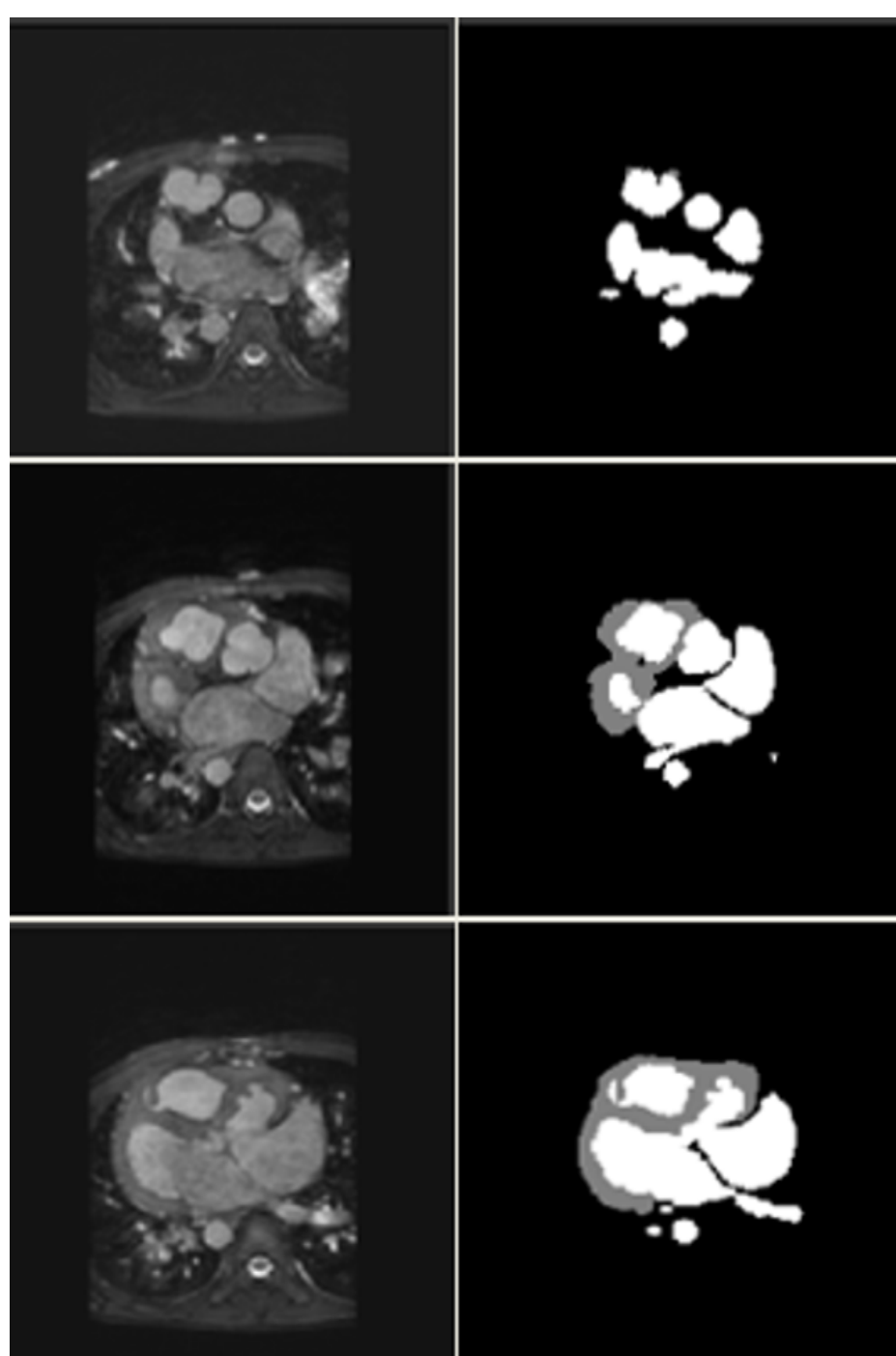


Fig4. Imagen y máscara de segmentado

Algoritmo de entrenamiento

El algoritmo de optimización empleado para el aprendizaje fue Adam (Adaptive moments), como función de costo se tomó la suma de la entropía categórica cruzada y el error en el coeficiente de dice, y como proceso de regularización se aplicó 'spatial dropout' en cada operación de convolución, el cual consiste no en desactivar unidades individuales de la red, si no en desactivar canales completos en forma aleatoria. El entrenamiento se llevó a cabo en una unidad de GPU, empleada en la plataforma de Google Colab. La evaluación del desempeño final en un conjunto de prueba fue llevada a cabo empleando el coeficiente de dice, el cual se calcula tomando en cuenta las tres categorías: fondo, volumen sanguíneo, paredes ventriculares.

Índice de Dice			
	Entrenamiento	Validación	Evaluación
mIoU	0.982	0.986	0.973

Tabla1. Resultados del valor promedio del coeficiente de Dice en los diferentes conjuntos.

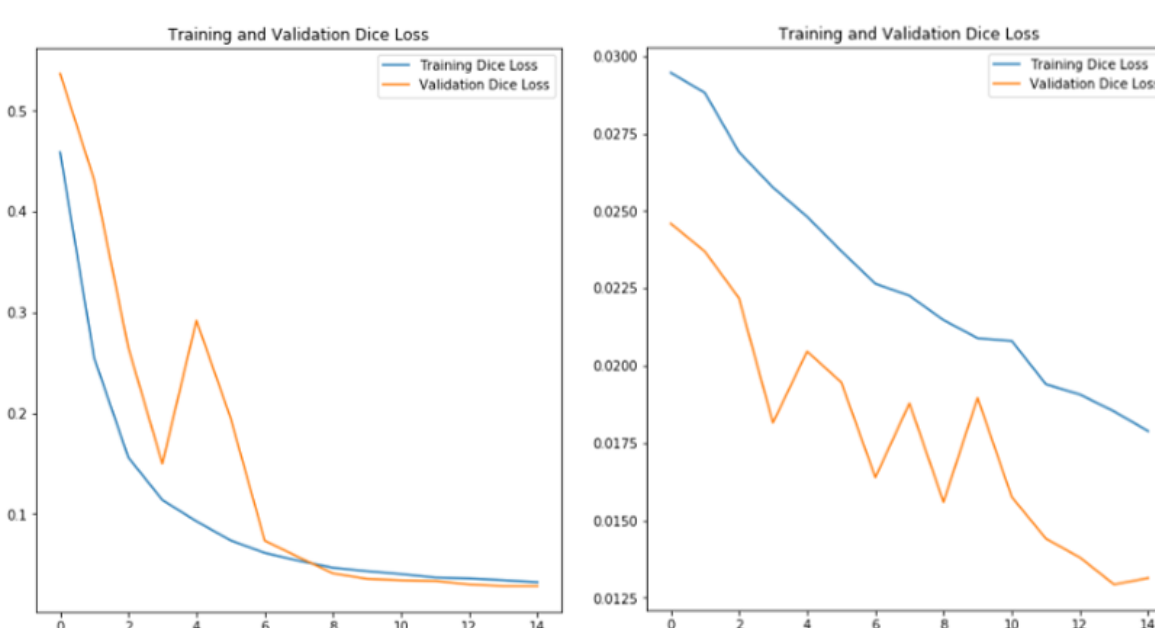


Fig5. Se grafica la función de costo. Izquierda) Primeras 15 épocas, Derecha) Últimas 15 épocas

Resultados del entrenamiento

Como forma de comparación cualitativa, se sobrepuso la máscara original y la máscara obtenida por nuestra red, con una copia de la imagen del corazón como se muestra a continuación.

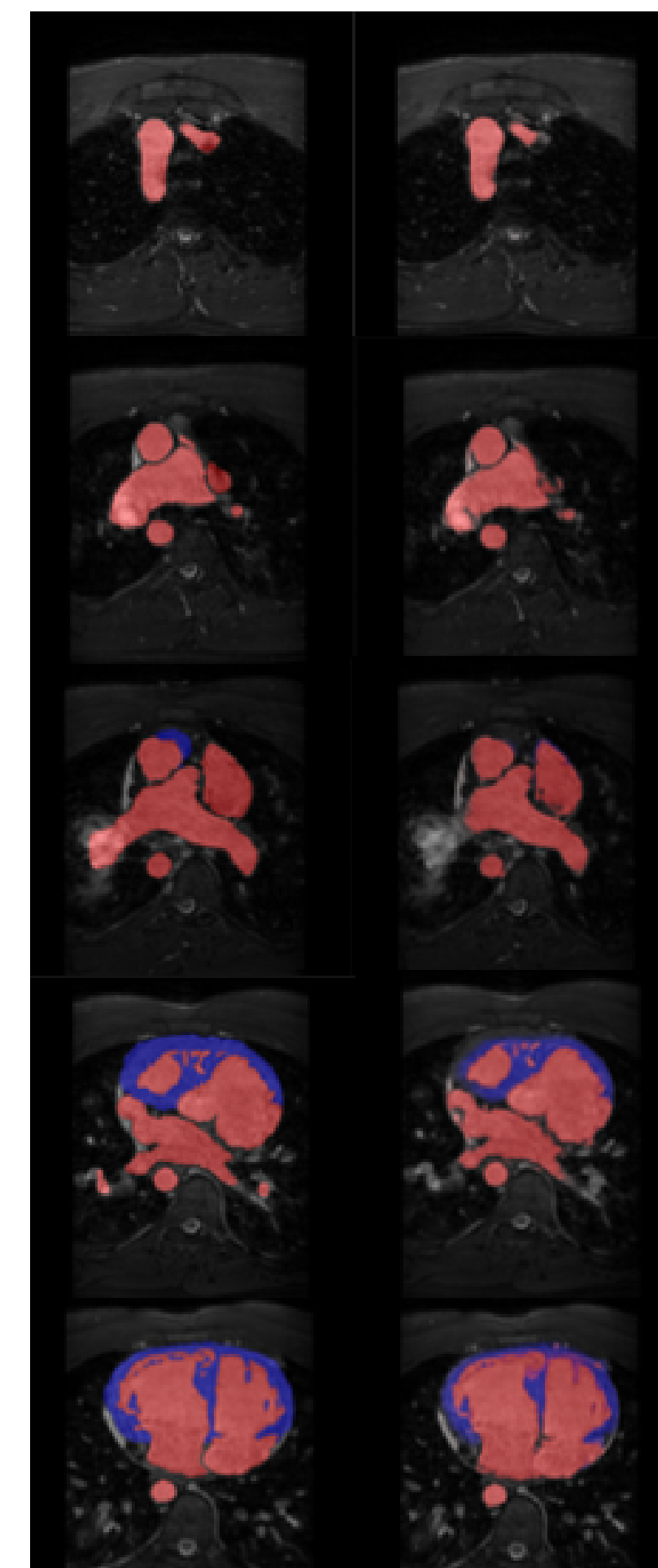


Fig6. Resultado de segmentación. Columna izquierda) Máscara de segmentación original sobrepuesta en la imagen médica. Columna Derecha) Máscara de segmentación que resulta de la red desarrollada, sobrepuesta en la imagen médica.

Conclusión

Fue posible demostrar que la arquitectura U-NET presenta una buena opción para llevar a cabo la segmentación de imágenes médicas, además de que modificaciones propuestas, como la convolución separable, así como adiciones, como el 'spatial dropout' presentan opciones que permiten alcanzar un buen rendimiento durante el proceso de aprendizaje de una red, así como lograr un buen desempeño al llevar a cabo esta tarea en particular, al mismo tiempo que se reducen el coste computacional y el número de parámetros requeridos.

Agradecimientos

Los autores agradecen el apoyo de los proyectos UNAM PA-PIIT IA103119 y IN116917.

Referencias

- [1] François Chollet. Xception: Deep learning with depthwise separable convolutions. CVPR, 2017.
- [2] Ian J. Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, USA, 2016. <http://www.deeplearningbook.org>.
- [3] Geva T. Powell A.J. Moghari M.H. Golland P. Pace D.F., Dalca A.V. Interactive whole-heart segmentation in congenital heart disease. In *Lecture Notes in Computer Science*, volume 9351, pages 80–88. Springer, Cham, 2015.
- [4] Brox T. Ronneberger O., Fischer P. U-net: Convolutional networks for biomedical image segmentation. IN: MICCAI 2015.
- [5] Christian Szegedy Sergey Ioffe. Batch normalization: Accelerating deep network training by reducing internal covariate shift. ICML, 2015.
- [6] Laurent Sifre. *Rigid-Motion Scattering For Image Classification*. PhD thesis, Ecole Polytechnique, CMAP, 2014.